DOCUMENT RESUME

ED 204 391 TM 810 386

TITLE Staff Report on the Federal Trade Commission Investigation of Coaching for Standardized Admission Tests.
INSTITUTION Federal Trade Commission, Boston, Mass.
PUB DATE [81]
NOTE 60p.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Achievement Rating; College Admission; *College Entrance Examinations; *Standardized Tests; *Test Coaching
IDENTIFIERS *College Entrance Examination Board; *Scholastic Aptitude Test

ABSTRACT

Whether coaching for standardized admission tests such as the Scholastic Aptitude Test (SAT) is effective was investigated. The Bureau of Consumer Protection of the Federal Trade Commission evaluated the effectiveness of two commercial coaching schools that offered preparation courses for the SAT. The researchers concluded that one of the two schools raised its students' scores an average of approximately 25 points each on both verbal and mathematical sections of the test. Staff found that a number of other studies, some funded by the College Board (CB) which sponsors the SAT, has also shown that students could make significant gains on the SAT as a result of participating in a coaching program. Staff also reviewed the descriptive materials regarding the SAT which are provided to students and schools each year by Educational Testing Service and CB. Staff concluded that although the harsh criticism of the CB coaching that had been levied in the late sixties had been modified in more recent materials, even the 1979-80 materials did not appear to recognize the possibility of meaningful score gains through coaching. The CB has indicated its intention to revise descriptive materials in the future. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

STAFF REPORT ON THE FEDERAL TRADE COMMISSION INVESTIGATION OF

COACHING FOR STANDARDIZED ADMISSION TESTS

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION

Tm 810 380

THE VIEWS CONTAINED IN THIS DOCUMENT ARE THOSE OF THE STAFF OF THE FEDERAL TRADE COMMISSION. THE VIEWS HAVE NOT BEEN FORMALLY ADOPTED BY THE COMMISSION.

BOSTON REGIONAL OFFICE
FEDERAL TRADE COMMISSION

STAFF REPORT

ON THE

FEDERAL TRADE COMMISSION INVESTIGATION

OF

COACHING FOR STANDARDIZED ADMISSION TESTS

The views contained in this document are those of the staff of the Federal Trade Commission. The views have not been formally adopted by the Commission.

EXECUTIVE SUMMARY

Standardized admission tests have recently become the subject of public debate. Critics have questioned whether the tests can accurately predict future academic achievement and whether the tests unfairly discriminate against the poor, minorities and culturally disadvantaged students. Others have also questioned the use of multiple-choice tests in making admissions decisions because such tests cannot reflect a student's creativity or motivation.

The Boston Regional Office of the Federal Trade Commission has been investigating one important issue regarding standardized admission tests - whether coaching for such tests is effective. The investigation was spawned by a concern that commercial coaching schools were making unfair and deceptive claims regarding effectiveness. These schools claimed that their programs could help students improve their scores on a variety of standardized tests by as much as one hundred points. At the same time, however, the Educational Testing Service ("ETS"), the largest of the testmakers, maintained that coaching would do little to help.

As part of the investigation, the Bureau of Consumer Protection of the Federal Trade Commission evaluated the effectiveness of two commercial coaching schools that offered preparation courses for the Scholastic Aptitude Test ("SAT"). The researchers concluded that one of the two schools was ef-

fective in raising its students' scores an average of approximately twenty-five points each on both the verbal and mathematical sections of the test. These findings led the staff of the Boston Regional Office to review other research on coaching for the SAT. Staff found that a number of other studies -some funded by the sponsor of the SAT itself, the College Board - had also shown that students could make significant gains on the SAT as a result of participating in a coaching program.

The question of whether coaching for the SAT is effective is an important one because of the role that the SAT plays in college admissions decisions. In a recent survey, 42% of the private four year colleges and 37.9% of the public four year colleges responding indicated that they had minimum SAT score requirements below which applicants generally are not considered eligible for admission. In light of the emphasis that students, high schools, colleges and even communities place on the results of the SAT, Boston Regional Office staff reviewed the descriptive materials regarding the SAT which are provided to students and schools each year by ETS and College Board. Staff wanted to determine whether the materials fully and accurately described the possible benefits of coaching. As a result of that review and discussions with ETS and College Board, the Boston Regional Office concluded that there was cause for concern. Though the harsh criticism of coaching

that had been levied in the late sixties had been modified in more recent materials, even the 1979-80 materials did not appear to recognize the possibility of meaningful score gains through coaching.

The Commission shared the Boston Regional Office's concerns about the 1979-80 materials and offered both ETS and College Board an opportunity to respond to those concerns. ETS submitted a lengthy report on coaching which included two statistical reanalyses of the Bureau of Consumer Protection's study. The ETS reanalyses largely confirmed the findings made by the Bureau.

The College Board submitted the descriptive materials on the SAT prepared for the 1980-81 testing year. The materials included a one page message on coaching sent to high schools which, for the first time, provides a more comprehensive explanation of the possible benefits of coaching. The message recognizes that there are different types of coaching programs and acknowledges that different results may occur. The message explains that some programs had been found to produce gains of about 10 points on the verbal section and 15 points on the mathematical section, while others had resulted in increases as high as 25-30 points on a section of the SAT. While the message may not reflect the still greater increases found in some studies of coaching, it represents an important move toward full disclosure. The other more

traditional materials on the SAT, which include a guide and bulletin for students and a guide for high school counselors and admissions officers, do not parallel the new message. These materials appear to recognize the findings made in a limited number of studies and acknowledge only the 10 and 15 point gains from coaching. The materials do not reflect the fact that some types of coaching can result in greater gains. By letter to staff of the Commission, College Board has indicated its intention to revise such descriptive materials in the future.

A finding that coaching can be effective for a standardized admission test such as the SAT has far-reaching educational implications. Questions are raised about the meaning of an examination designed to measure verbal and mathematical abilities said to be developed over a lifetime of learning if scores can be meaningfully changed in a few weeks or months. Questions of fairness are also raised if coaching can have a substantial impact on scores but such coaching is not equally available to all students, particularly the less affluent. Much research remains to be done before these questions can begin to be addressed. More must be known about what precisely the SAT measures and why coaching can help students achieve meaningful score gains. More must be known about the availability of coaching. More must be learned about the kinds of students who may benefit

the most from a coaching program and what impact unequal access to coaching may have on students' access to higher education.

As a law enforcement agency charged with protecting the public from unfair and deceptive practices, the Federal Trade Commission has concentrated its efforts on assuring that organizations involved in the standardized testing industry make full and fair disclosure regarding the coaching issue. The educational policy questions raised by the Federal Trade Commission's investigation into the coaching issue are ones which should be addressed by those with the mandate and greater expertise in this important area.

I. Introduction

Standardized admission tests have recently become the focus of much attention. The validity and value of such tests has been the subject of public debate. Congress, as well as many state legislatures, has been considering disclosure legislation regarding these tests. The Federal Trade Commission has conducted an investigation into one important issue regarding standardized admission tests -the effectiveness of special preparation for such tests. Because of the role that standardized admission tests play in the educational lives of young people and in light of the serious educational policy questions raised by the coaching issue, the Boston Regional Office believes that it is important to fully report on the Federal Trade Commission's investigation. The views contained in this document are those of the staff of the Federal Trade Commission. The views have not been formally adopted by the Commission.

In late 1976 the Boston Regional Office of the Federal Trade Commission was authorized to conduct an industry-wide investigation into the validity of claims made regarding the effectiveness of coaching for standardized admission tests. The investigation was spawned by concern that commercial coaching schools were misleading prospective students by unfair and deceptive claims of effectiveness. The desire to assess those claims resulted in the undertaking of a

detailed statistical analysis of the effectiveness of commercial coaching school programs by the Federal Trade Commission's Bureau of Consumer Protection. That study concluded that for at least some students coaching for the Scholastic Aptitude Test could be effective. A review of other research on the effectiveness of coaching for the Scholastic Aptitude Test made clear that the Bureau's positive findings were not an isolated occurrence. Other studies had reached similar conclusions about the effectiveness of coaching.

The Bureau's positive findings raised concern as to whether the possible effects of coaching for the Scholastic Aptitude Test were properly being reflected in the materials which the testmaker, Educational Testing Service, and the test sponsor, College Board, provided to students and educators. The Boston Regional Office's review of those materials revealed that intense criticism of coaching had been levied in the materials distributed in the late sixties. While that criticism was tempered over the ensuing years, the overall message remained that coaching for the Scholastic Aptitude Test was unlikely to be of any significant benefit. A 1980 message on coaching sent to secondary schools by the College Board provides, for the first time, a more comprehensive discussion of the possible benefits of various forms or coaching for the Scholastic Aptitude Test. Similar disclosures have not yet been made in other

student and educator descriptive materials on the Scholastic Aptitude Test, although the College Board has indicated that 1981-82 materials will contain revised information about coaching.

A finding that coaching can be effective for a standardized admission test such as the Scholastic Aptitude Test raises a number of very important educational policy questions. Though the Federal Trade Commission has committed its resources to assuring that organizations involved in the standardized testing industry make full and fair disclosure regarding the coaching issue, it has neither the mandate nor the expertise to deal with all of the important educational questions. Staff has included a discussion of some of those questions in this report so that the expertise and resources of the appropriate government agencies, as well as private researchers, may be directed to finding the answers.

II. Background on the Standardized Admission Test Industry

A. Standardized Admission Tests

Standardized admission tests are professionally developed examinations, often administered nationwide and most often in multiple-choice format, which are used in the evaluation of applicants for admission to a wide variety of educational institutions, including private secondary schools, colleges and graduate schools. Each year, over a million students take some form of standardized admission test.

There are a number of both profit and nonprofit corporations engaged in the development of standardized admission tests. The tests involved in our investigation were developed by the Educational Testing Service ("ETS"). ETS, a nonprofit corporation located in Princeton, New Jersey, with reported revenues of over ninety million dollars for its 1979 fiscal year, [1] is the largest testmaker and administrator of standardized admission tests in the nation. Its tests include the Secondary School Admission Test, used for grades six through eleven, the Scholastic Aptitude Test ("SAT"), used by colleges and universities throughout the country, the Law School Admission Test ("LSAT"), used in law school admissions, and the Graduate Record Examinations, used by graduate schools of various types. Organizations of educators and educational institutions sponsor each of these tests. For example, the College Board, a nonprofit membership organization composed of colleges, schools, school systems and education associations, sponsors the SAT and the Law School Admission Council, a nonprofit membership organization of law school representatives, sponsors the LSAT.

---

1 1979 Annual Report 22 (Educational Testing Service 1980).

B. Coaching for Standardized Admission Tests

Special preparation for standardized admission tests, often referred to as "coaching," can take many forms. Coaching can be as simple as a few hours of drill on test questions patterned after those expected to be on the standardized admission test. It can also be much more intensive and complex, including review of the substantive content to be tested by the exam, as well as instuction on test-taking strategies and practice on sample tests. Coaching programs are offered by a variety of organizations, including commercial coaching schools, community groups, private tutors, and, in some cases, by secondary schools and colleges.

The Boston Regional Office's initial investigation into the coaching issue found that commercial coaching courses were being offered in every state and in virtually every major city or college campus throughout the country. Tuition costs for such courses have ranged from less than $50 to at least $300. Some of these schools provide no more than a weekend of instruction, while others provide over forty hours of in-class time, as well as homework materials and practice tapes.

III. Importance of the Coaching Issue

The question of whether coaching for standardized admission tests is effective is important both to the students whose abilities are being evaluated by the tests and to

the schools that use the tests as at least one basis by which to evaluate applicants for admission. Students concerned with maximizing their chances for admission to the schools of their choice want to know whether time and money would be well spent on special preparation. Schools need to know as much as possible about the coaching issue in considering what weight they will give to standardized admission tests in making admissions and placement decisions. As a practical matter, the more extensively schools rely on standardized test scores in making admissions decisions, the more important the question of coaching's effectiveness becomes.

Because the major focus of the investigation has been on coaching for the SAT, it is important to discuss more fully here the role that the SAT plays in college admissions decisions. The SAT is taken by over a million students each year. It is part of the application process at hundreds of private and public colleges and universities throughout the country.[2] The SAT consists of multiple-choice questions, with separate mathematical and verbal scores reported on a scale of 200 to 800.[3] For students in the 1978 graduating

---

2 Peterson's Annual Guide to Undergraduate Study: 1980 Edition 37-59 (J.Hunter ed. 1979) [hereinafter cited as Peterson's Guide].

3 Educational Testing Service, ATP Guide for High Schools and Colleges 1979-81, at 4, 8 (College Entrance Examination Board 1980).

class, the fiftieth percentile rank was a verbal score of 430 and a mathematical score of 470.[4] ETS describes the SAT as measuring "developed verbal and mathematical reasoning abilities that are related to successful performance in college."[5] Its purpose is said to be to serve as a predictor of academic performance in college.[6]

The precise role which the SAT plays in the college admissions process is currently the subject of some controversy. Critics of the test are concerned that it plays too important a role in those decisions.[7] In responding to such concerns,

---

4 Id. at 15.

5 Id. at 4.

6 Id. at 21. The degree to which the SAT accurately predicts college performance has recently been the subject of heated debate. Two recent publications strenuously challenged the predictive validity of the test. W. Slack & D. Porter, "The Scholastic Aptitude Test: A Critical Appraisal," 50 Harv. Ed. Rev. 154, at 164-69 (1980); A. Nairn & Associates, The Reign of ETS: The Corporation that Makes Up Minds 55-160 (1980). Rex Jackson of ETS responded to the Slack and Porter article in, "The Scholastic Aptitude Test: A Response to Slack and Porter's 'Critical Appraisal'," 50 Harv. Ed. Rev. 382 (1980), to which Slack and Porter responded in "Training, Validity, and the Issue of Aptitude: A Reply to Jackson," 50 Harv. Ed. Rev. 392 (1980). ETS responded to Nairn's criticism in two February 1980 pamphlets, Test Use and Validity and Test Scores and Family Income.

7 See, e.g., A. Nairn & Associates, The Reign of ETS: The Corporation that Makes Up Minds, 5-14, 47-50 (1980). Dean Vito Perrone of the Center for Teaching and Learning at the University of North Dakota expressed concern about undue reliance on the test because of the impact

(Footnote Continued)

ETS has noted that admissions officers have been advised by ETS and College Board that test scores should not be the sole factor used in making an admissions decision and that recent data indicates that colleges today are simply not very selective, that most admit a large proportion of their applicants.[8]

A 1979 survey conducted by the College Board and the American Association of Collegiate Registrars and Admissions Officers sheds some light on the issue. Of the colleges responding which indicated that they did not have open-door policies, only 1.8% said that admissions test scores were the single most important factor in making admissions

---

7 (Footnote Continued)

on poor and minority students. Truth in Testing Act of 1979; The Educational Testing Act of 1979: Hearings on H.R. 3564 and H.R. 4949 before the Subcomm. on Elementary, Secondary, and Vocational Education of the House Committee on Education and Labor, 96th Cong., 1st Sess. 455-59 (1979) (Appendix to the prepared statement of Vito Perrone, Dean, Center for Teaching and Learning, University of North Dakota). [Hereinafter references to the record in the hearings are cited as Hearings on Testing.] Paul Pottinger, Executive Director of the National Center for the Study of Professions, expressed concern about reliance on standardized admission tests in general: "Multiple-choice aptitude tests are unfair because they fail to assess the more critical aspects of competence that determine how effective students could be as workers and productive citizens." Hearings on Testing, supra at 475.

8 Educational Testing Service, Test Use and Validity 11-12 (February 1980).

decisions. However, 56.6% of those colleges cited admissions test scores as a very important factor and 33.1% cited them as one of several factors.[9] Perhaps more importantly, 42% of the private four year colleges and 37.9% of the public four year colleges responding indicated that they had minimum SAT scores below which applicants generally are not considered eligible for admission.[10]

Other sources also indicate that a difference in a very small number of points on the SAT may, in some cases, be very important. For example, College Board's The College Handbook 1980-81 noted the "Basis for selection" at Florida State University was that in-state students must have a 2.0 academic school grade average and a combined score of 80[illegible] on the SAT for applicants who had taken that standardized admission test.[11] The numbers for out-of-state applicants were an average of approximately 3.0 and a combined score of 1000.[12] At Arizona State University the "Basis for selection" was said to

---

9 College Board, November 1, 1979 Press Release, Table 1 a.

10 "An Overview of Findings from the College Board-AACRAO Survey of Undergraduate Admissions Policies, Practices, and Procedures" 12 (Prepared for the 1979 College Board Annual Meeting, New Orleans, Louisiana, October 30, 1979).

11 The College Handbook 1980-81, at 261-62 (M.Matheson ed., College Entrance Examination Board 1980).

12 Id.

be rank in top 50% of graduating class or a minimum SAT combined score of 930 for in-state applicants and 1010 for out-of-state applicants who take that exam.[13] An even more complex use of minimum SAT scores is noted in the 1981-82 undergraduate application materials for the University of California. The materials include what is termed the "University of California Freshman Eligibility Index" for in-state applicants. That index lists the required grade point average in certain specified courses along with the required scores on standardized admission tests. For example, if the applicant has a 2.79 grade point average, the required SAT combined score is 1580. For a 3.01 grade point average, the corresponding SAT score is 1060. California residents with grade point averages of 3.3 or above are eligible for admission regardless of any test scores.[14]

---

13 Id. at 34.

14 University of California Undergraduate Application Packet 1981-82, at 13, 15. Non-residents must have grade point averages of 3.4 or higher in the required secondary school courses. Id. at 16. The materials note that the University recognizes that some students have not had the same opportunities to prepare for college work and the University offers special admissions consideration for those who can offer evidence of their ability to achieve even though they have not met traditional requirements. The materials note that such applications are encouraged even though a limited member of persons may be admitted in that category. Id. 18.

In assessing the role of SAT scores in college admissions, it is also important to consider the degree to which college bound students rely on their SAT scores in selecting the schools to which they will apply. The answer to this question is not clear.[15] A 1978 survey of SAT takers notes that 16% of the students responding indicated that their test scores had caused them to change their minds about the kinds of schools they were considering, in most cases resulting in their lowering their sights.[16] Students clearly do have information at hand with which to engage in self-selection on the basis of SAT scores.

The score report sent to each SAT taker explains his or her score and locates it relative to the scores of other test takers.[17] In addition to whatever information on SAT scores the colleges provide in their catalogues, college reference guides include detailed information about the SAT scores of enrolled freshmen whenever it is available.[18]

---

15 R. Hartnett & R. Feldmesser, "College Admissions Testing and the Myth of Selectivity: Unresolved Questions and Needed Research," American Association of Higher Education Bulletin, March 1980, at 5-6.

16 Response Analysis Corporation, "SAT Monitor Program" 47 (July 1978).

17 Educational Testing Service, ATP Guide for High Schools and Colleges 1979-81, at 13-17 (College Entrance Examination Board 1980).

18 Peterson's Guide, supra n.2; J. Cass & M. Birnbaum, (Footnote Continued)

By a review of such reference guides, prospective applicants can learn that at the California Institute of Technology one hundred percent of 1978-79 freshmen mathematical scores on the SAT were 600 or over.[19] Over ninety-five percent of the 1978-79 freshmen SAT mathematical scores were 500 or over at the Rose-Hulman Institute of Technology,[20] Rice University,[21] Stanford University,[22] and the Colorado School of Mines.[23] At Yale University,[24] Haverford College,[25] Reed College[26] and Barnard College[27] over ninety percent of 1978-79 freshmen verbal scores were 500 or above.

The SAT is clearly viewed as an important part of the educational system. Students do take their individual

---

18 (Footnote Continued)

Comparative Guide to American Colleges (9th ed. 1979).

19 Peterson's Guide, supra n.2, at 40.

20 Id. at 53.

21 Id. at 52.

22 Id. at 55.

23 Id. at 42.

24 Id. at 59.

25 Id. at 45.

26 Id. at 52.

27 Id. at 39.

SAT scores seriously.[28] In fact, concern has been raised that too many students may erroneously gauge their self-worth or mental capacities by their SAT scores.[29] Secondary schools are very concerned about the average SAT scores of their students.[30] Competitive colleges tend to use average SAT scores to evaluate the quality of their freshmen classes.[31] Communities view SAT scores as at least one

---

28 See R. Moll, Playing the Private College Admissions Game 144-46 (1979) (Moll, who has served as admissions director at Bowdoin and Vassar colleges, recounts an instance where a high school counselor introduced a group of candidates with a list that contained the students' names and SAT scores. "No wonder candidates feel their scores are tattooed on their foreheads and it is their fate to wander through life muttering at every turn, 'I am a 510.'"); J. Fallows, "The Tests and the 'Brightest'," The Atlantic, Feb. 1980, at 44. ("I have yet to meet a high school student who did not take the tests as a measure of how 'smart' he was.") An excerpt from the July 18, 1979 Chicago Sun-Times included in the record of the truth in testing hearings is also of note. "College entrance exams have replaced the draft as perhaps the most important and mysterious rite of passage for young people." Hearings on Testing, supra n.7, at 666.

29 W. Slack & D. Porter, "Training, Validity, and the Issue of Aptitude: A Reply to Jackson," 50 Harv. Ed. Rev. 392, at 399 (1980). The authors are concerned by the use of the word "aptitude" in the name of the SAT. "[S]tudents who have interpreted low SAT scores as a reflection of their aptitude have been seriously wronged; their capacity to learn has been disparaged and their self-esteem endangered."

30 See S. Thomson & N. DeLeonibus, Guidelines for Improving SAT Scores (National Association of Secondary School Principals 1978).

31 R. Moll, Playing the Private College Admissions Game, 146-47 (1979).

indication of the quality of their school systems.[32] Nationally, as an ETS researcher has noted, "the average test score has become a kind of educational Dow-Jones index."[33]

IV. Federal Trade Commission Investigation of the Effect of Coaching on Standardized Admission Tests

A. Focus of the Investigation

The Boston Regional Office ("BRO") began its investigation into the coaching issue when staff became aware of commercial coaching school advertising claims regarding the effectiveness of their programs in preparing students to take a variety of standardized admission tests. These schools claimed to be able to improve students' scores by as much as 100 points. Yet, ETS had maintained that coaching would do little to improve students' scores. BRO was concerned that students interested in improving their chances for admission to the college or graduate school of their choice were wasting considerable time and money in reliance on the coaching schools' claims of effectiveness. Pursuant to Section 5 of the Federal Trade Commission Act, 15 U.S.C. § 45, the Commission authorized the

---

32 See The Boston Globe, June 30, 1978, at 1, col. 2.

33 Donlon, T., "The SAT: Past and Present," The College Board Review, Fall 1979, at 29.

BRO to evaluate whether the commercial coaching schools' claims were unfair or deceptive.

BRO specifically examined twenty-two commercial coaching schools, focusing on those schools that offered preparation for the SAT and/or the LSAT. Staff evaluated voluminous documents submitted by the schools to substantiate their claims of effectiveness. In addition, in order to have an independent assessment of the effectiveness of coaching, the BRO undertook a statistical analysis of the effectiveness of commercial coaching for both the SAT and the LSAT.

B. Findings of the FTC Regarding the Effectiveness of Coaching

The BRO analyzed an existing situation, comparing the scores of students who had voluntarily enrolled in commercial coaching courses with a group of control students who had chosen not to enroll in those courses. Thus, the BRO study was not based on an experimental design; students were not randomly assigned to coached and control groups. The coached sample was developed from the enrollment lists submitted by the commercial coaching schools. The control sample was developed from ETS records of students who lived in the same geographical area and who had taken the same administration(s) of the SAT or LSAT as the coached students, but who were not identified as coaching school enrollees.

In September of 1978, the BRO submitted its memorandum on the investigation to the Federal Trade Commission's Bureau of

Consumer Protection ("BCP") in Washington, D.C. In its memorandum, BRO concluded that its statistical analysis showed that coaching was dramatically effective for the SAT.[34] While the LSAT was seen to be susceptible to coaching in a general sense, the results showed areas where the effects were marginal.[35]

The BCP reviewed the BRO's memorandum and found that the statistical methodology utilized did not sufficiently control for differences between the coached and control groups to permit the conclusions that had been reached. BCP undertook a refined statistical analysis of the data that had been collected. The focus of the reanalysis was to isolate the effect of coaching by controlling for a number of background differences between the coached and control groups and by analyzing the impact of self-selection, the fact that the coached students had voluntarily decided to attend a commercial coaching course. The BCP researchers concluded that the Boston data contained enough background information on the SAT test takers to perform a meaningful reanalysis, but that sufficent information was not available for the LSAT.

The BCP reanalysis of the SAT data found that coaching could have a significant positive impact on scores, though

---

34 Boston Regional Office, "Staff Memorandum on the Effects of Coaching on Standardized Admission Examinations" 157 (September 1978).

35 Id.

the point value of the score improvement found by the BCP was smaller than that found by the BRO. In the BCP report, the researchers concluded that coaching at one of the two schools evaluated was effective in raising SAT scores by an average of approximately 25 points each on the mathematical and verbal sections of the test.[36] The students at the effective school tended to be underachievers on standardized examinations, and thus the issue of whether other students might also benefit from coaching was not resolved in the report.[37] The very limited data sample that was available regarding students at the effective school who were not underachievers indicated the possibility that they might also benefit from coaching.[38]

---

36 Bureau of Consumer Protection, "Revised Statistical Analyses of Data Gathered by Boston Regional Office of the Federal Trade Commission" Executive Summary (March 1979) [hereinafter cited as BCP Report]. Because of sample size limitations, the BCP analyzed data from only two schools.

Students who take the SAT for a second time are expected to experience score increases on average as a result of practice and growth. When references are made in this report to score increases attributable to coaching, those increases are above and beyond what would otherwise be expected from such practice and growth.

37 Id. As used in the BCP Report, "underachiever" means a student who scores lower on the standardized test than would have been predicted from socio-economic and demographic background data.

38 Id.

The Commission recognized the importance of having the studies in the public domain and released both the BRO and BCP statistical analyses in May of 1979. Underlying data for the studies, contained on computer tapes, were also made available through the National Archives.

C. The Significance of the FTC's Findings in Light of Prior Research

In order to assess the significance of the conclusions reached in the BCP report, BRO analyzed other research in the area. While each different study on the effectiveness of coaching, including the BCP study, must be evaluated in terms of its own statistical design and limitations, a number of studies in addition to the BCP study do point to the effectiveness of coaching for the SAT.

Many of the earliest studies on coaching for the SAT were summarized by College Board in a booklet entitled Effects of Coaching on Scholastic Aptitude Test Scores, copyrighted in 1965 and again in 1968. The booklet explained that the word "coaching" was used to refer "to a variety of methods used in attempting to increase in a relatively short time students' mastery of the particular skills, concepts, and reasoning abilities tested by the SAT."[39] The booklet

---

39 College Entrance Examination Board, Effects of Coaching on Scholastic Aptitude Test Scores 4 (College Entrance Examination Board 1968) [hereinafter cited as Effects of Coaching].

included a lengthy statement by the College Board trustees issued as a result of the studies testifying that "increases in scores on the SAT that may result from coaching are negligible."[40]

A recent article in the Harvard Educational Review has questioned the accuracy and completeness of that booklet.[41] The authors conclude that "there is ample evidence that students can successfully train for the SAT," and that such evidence was available at the time the College Board booklet was published in 1968.[42]

---

40 Id. at 8.

41 W. Slack & D. Porter, "The Scholastic Aptitude Test: A Critical Appraisal," 50 Harv. Ed. Rev. 154, at 155-64 (1980). The authors assert that College Board failed to reference studies by Pallone (1961) and Marron (1965). Id. at 158-59. The Marron studies involved high school graduates who spent seven months in private preparatory schools. The studies did not include control groups; Slack and Porter estimate the mean gains to have been 41.1 points on verbal scores and 67.6 points on mathematical scores over and above the gains that would have been expected from practice and growth. Id. at 159. The Pallone study involved one program of 20 students in a six-week, 90 minutes per day program of verbal instruction and a 50 minutes per day verbal program for 100 students from September to March. Again no control groups were included. Slack and Porter estimate the mean gains as 84.7 points on verbal for the six-week program and 79.1 points on verbal for the longer course. Id. at 158-59. The six-week Pallone study is discussed infra at 23-24.

42 Id. at 164. As noted supra at n.6, ETS has responded to the Slack and Porter article. R. Jackson, "The Scholastic Aptitude Test: A Response to Slack and
(Footnote Continued)

Looking exclusively at the studies summarized by College Board, one finds that for some groups of students and some forms of preparation score improvements of more than 15 points were reported. In one study, which involved tutoring in groups of two or three, coached students showed average gains on the mathematical section of the SAT of about 21 points after approximately 18 hours of coaching and about 26 points after 36 hours.[43] In another study, one group of students showed average gains of 18 points on the verbal section after preparation on special exercises developed by ETS to resemble items on the test.[44] Significant gains on the mathematical section were also reported in a third

---

42 (Footnote Continued)

Porter's 'Critical Appraisal'," 50 Harv. Ed. Rev. 382 (1980). As part of that response, Jackson asserts that Slack and Porter failed "to draw clearly the critical distinction between programs of short-term drill and practice designed to yield quick increases in scores, and longer-term educational programs designed to have lasting effects." Id. at 383. Slack and Porter responded to the Jackson article in "Training, Validity, and the Issue of Aptitude: A Reply to Jackson," 50 Harv. Ed. Rev. 392 (1980).

43 Effects of Coaching, supra n.39, at 18-21. The 18 and 36 hours of coaching included instruction directed at both the mathematical and verbal sections of the test.

44 Id. at 14-17.

study for students not taking senior math, an average gain of about 29 points over the control group.[45]

Moreover, the studies summarized by College Board may not all be relevant to an analysis of more sophisticated coaching courses offered to today's SAT applicants. A 1972 study sponsored by College Board noted two problems with earlier studies in discussing why further research on coaching for the mathematical section of the SAT was needed. First, the instruction provided in the majority of the studies was, "where its nature can be ascertained, rather scanty. There appeared to be little or no systematic attempt to identify the information and skills needed to perform well on the test and to develop materials to meet these needs."[46] Second, most of the prior research involved students at the extremes of the ability range rather than the more heterogeneous group of students who now take the SAT.[47]

The 1972 study itself involved seven week, 21 hour, coaching programs for high school juniors at twelve schools.[48] Detailed lesson outlines, and other instructional material

---

45 Id. at 12-14.

46 L. Pike & F. Evans, Effects of Special Instruction for Three Kinds of Mathematics Aptitude Items 4 (College Entrance Examination Board 1972).

47 Id.

48 Id. at Abstract.

developed by ETS, were provided to the instructors who were trained in a two-day workshop.[49] The authors concluded that students could be prepared with respect to each of the three types of math item formats evaluated in the study.

> Results of the statistical analysis showed that each of the three item formats was susceptible to the special instruction specifically directed toward it. The complex or novel item formats appeared to be more susceptible than the relatively straightforward item format. Female volunteers were found to be slightly less able mathematically at the outset, and to benefit somewhat less from the instruction than male volunteers. Mean gains of nearly a full standard deviation obtained by the groups instructed for the complex or novel formats were considered to be of practical consequence and likely to influence admissions decisions.[50]

Significantly, the results of the study were consistent for all twelve schools.[51] One of the two authors, Lewis Pike, later explained in a 1978 literature review of coaching studies, that a judicious combination of instruction for the two item formats used at the time the 1972 study was conducted "would be expected to yield an STI [short term instruction] effect of about 33 points."[52]

---

49 Id. at 13 and 10.

50 Id. at Abstract.

51 Id.

52 L. Pike, Short-term Instruction, Testwiseness, and the Scholastic Aptitude Test: A Literature Review With Research Recommendations 16 (College Entrance Examination Board Research and Development Reports, 77-78, No. 2, January 1978) [hereinafter cited as Literature Review].

In the 1978 literature review, which was published as a College Board research and development report, Pike discussed and analyzed both negative and positive findings in prior studies of short term instruction ("STI") for the mathematical ("SAT-M") and verbal ("SAT-V") sections of the SAT. The abstract for the review noted that discrepancies "were generally resolved in favor of recognizing meaningful STI effects for the SAT-M, but remain unresolved for the SAT-V."[53] Studies discussed showing positive results for preparation relevant to the verbal section did point to at least the possibility of meaningful verbal gains. In one study, conducted by Pallone, 20 students participated in a focused program of instruction on reading, vocabulary and logical reasoning abilities for 90 minutes a day over six weeks.[54] Those students experienced an average gain of 98 points on the verbal section of the SAT.[55] Because the study did not include a control group, Pike estimated the effect of coaching by subtracting the gain one would have expected from growth and practice. Pike estimated that the instruction had produced approximately 60 point gains.[56]

---

53 Id. at i.

54 Id. at 19.

55 Id.

56 Id. Pike used "the gains experienced by controls at
(Footnote Continued)

In discussing the negative and positive findings of the verbal studies reviewed, Pike noted that the most marked difference between the Pallone study and those with negative findings was the degree to which instruction in Pallone's study was subtantive and controlled.[57] Pike also noted some general considerations important to evaluating mixed research findings. Because of the importance of such considerations to an understanding of the research on coaching, they are quoted here.

> Note that, in principle, a single study showing substantial positive gains cannot be countered or refuted by any number of studies failing to get positive results. The only near-exception would occur in the event of a well-designed replication study that failed to show similarly positive results. In that case, there would be a discrepancy needing further study and resolution. Similarly, it would be fallacious to infer, from mixed results across studies on a topic such as STI effects, that across-study inconsistencies justify the conclusion that there are no meaningful effects. Mixed results can mean, as exemplified in Jacobs' (1966) discussion of differences on ECT [English Composition Test] score changes from one experimental group to another, that an effort should be made to find out why instruction was effective in some places but not in others. This observation is particularly true when making comparisons between studies in which little account was taken of either examinee or instructional characteristics. A third observation is that there has been a considerable

---

56 (Footnote Continued)

the Bronx High School of Science as a rough (and probably conservatively high) estimation of control subject gains." Id.

57 Id. at 60-61.

> emphasis in most discussions of STI on the overall magnitude of its effects, with little consideration given, especially when stating final conclusions, to differences among examinees, among STI curricula, or among item formats and other item characteristics.[58]

A recent paper, based on research supported by the College Board, sheds further light on the possibility of meaningful improvements from coaching for the verbal section of the SAT. The researchers, Alderman and Powers, evaluated special verbal preparation programs being offered in eight secondary schools.[59] The instructional content of the various programs was not controlled by the researchers; the effect of existing school programs was measured.[60] Most of the programs were extracurricular and only one involved a program of over twelve hours.[61] "Most schools followed a commercial review book in their classes; one public school (school A) distributed teacher-made materials intended for use in tutoring individual students in a self-paced manner."[62] The estimated treatment effects on the verbal sec-

---

58 Id. at 35-36.

59 D. Alderman & D. Powers, The Effects of Special Preparation on SAT-Verbal Scores, Abstract (College Entrance Examination Board Research and Development Reports, 78-79, No. 4, February 1979).

60 Id.

61 Id. at Table 2.

62 Id. at 7.

tion of the SAT ranged from +28.39 points to -2.75 points.[63] The school at which the 28 point gain was found was the school that used the teacher-made materials described above.[64]

Thus, the BCP findings regarding the effectiveness of coaching for both the mathematical and verbal sections of the SAT are not an isolated instance of positive findings. Other research, particularly studies which focused on comprehensive and better structured programs, also found positive effects of coaching for the SAT.

D. SAT Descriptive Materials for 1979-80

In light of prior research and the BCP's findings regarding the effectiveness of coaching for the SAT, the staff of the BRO met with ETS and College Board and undertook a review of SAT descriptive materials regarding whether the possible benefits of coaching had been properly disclosed by ETS and College Board. The materials reviewed included pamphlets for both students and educators.

Perhaps the strongest criticism of coaching was an excerpt from the College Board's Effects of Coaching on Scholastic Aptitude Test Scores booklet included in pamphlets for counselors and admissions officers in 1965 through 1968.[65]

---

63 Id. at Table 5.

64 Id. at 7 and Table 5.

65 College Board Score Reports: A Guide for Counselors
(Footnote Continued)

The excerpt was the statement by the trustees of the College Board referred to at pages 18-19 supra.

The trustees stated that "The evidence collected leads us to conclude that intensive drill for the SAT, either on its verbal or its mathematical part, is at best likely to yield insignificant increases in scores. The magnitudes of the increases which have been found vary slightly from study to study, but they are always small and appear to be independent of the particular method of coaching used and of the level of ability of the students being coached."[66] They expressed their belief that coaching would not result in important score increases even if conducted under different circumstances and with different students from those in earlier studies.[67]

---

65 (Footnote Continued)

and Admissions Officers 52-54 (College Entrance Examination Board 1965) [hereinafter cited as 1965 Guide]; College Board Score Reports: A Guide for Counselors and Admissions Officers 51-54 (College Entrance Examination Board 1966) [hereinafter cited as 1966 Guide]; College Board Score Reports: A Guide for Counselors and Admissions Officers 56-58 (College Entrance Examination Board 1967) [hereinafter cited as 1967 Guide]; College Board Score Reports: A Guide for Counselors and Admissions Officers 57-59 (College Entrance Examination Board 1968) [hereinafter cited as 1968 Guide].

66 1965 Guide at 52; 1966 Guide at 52; 1967 Guide at 56; 1968 Guide at 57.

67 1965 Guide at 53; 1966 Guide at 52; 1967 Guide at 57; 1968 Guide at 58.

References to coaching were harsh. The trustees noted that parents had "demanded that the schools divert teaching energy and time to a kind of drill that is obnoxious to educators of every philosophy."[68] They closed their nine paragraph statement by noting that they were most concerned because they saw "the educational process unwillingly corrupted in some schools to gain ends which we believe to be not only unworthy but, ironically, unattainable."[69]

Though this harsh criticism of coaching was modified in later years, even the 1979-80 materials continued to indicate that special preparation was unlikely to be of significant benefit.

The student guide, Taking the SAT, simply explained that:

> The verbal and mathematical abilities measured by the SAT develop over years of study and practice. Drilling or last-minute cramming probably will not do much to prepare you for the test. However, if you are not taking a mathematics course, a review of mathematics concepts, such as those given in this booklet, will be useful.[70]

---

68 1965 Guide at 53; 1966 Guide at 52; 1967 Guide at 56; 1968 Guide at 58.

69 1965 Guide at 54; 1966 Guide at 53-54; 1967 Guide at 58; 1968 Guide at 59 (Footnote omitted in each reference.) The footnote explained that the statement had been prepared "for the information and advice of schools and colleges that had expressed concern about test coaching."

70 Educational Testing Service, Taking the SAT 3 (College Entrance Examination Board 1978).

Students were told that studying the sample questions, directions and explanations provided in the guide, which included a complete sample test, would help them prepare for the test.[71]

The Student Bulletin 1979-1980 addressed the coaching question in three sentences.

> Before the test date, read the appropriate descriptive booklet, Taking the SAT or About the Achievement Tests, which will give you a better understanding of the test. Studies have shown that cramming does not raise scores, but knowing what to expect can be helpful. A good night's sleep and a nourishing breakfast before you take a test are also helpful.[72]

The 1979-80 pamphlet for high school guidance counselors and college admissions officers, ATP Guide for High Schools and Colleges 1979-81, did include a lengthier discussion regarding special preparation. The educator pamphlet stated that:

> Over the past 25 years, the College Board has conducted many studies on the effect of special preparation programs on SAT score results and has supported the independent investigation of the topic by others. These studies consistently seem to demonstrate that "coaching," in the sense of intensive drill on sample test questions, does not lead to any significant improvement in students' scores. Special preparation of other kinds, however, may be helpful to some students.

---

71 Id.

72 Educational Testing Service, Student Bulletin 1979-1980, at 13 (College Entrance Examination Board 1979).

> It is especially important for all students to be familiar with the various types of questions in advance of the test. Students should know what the test is about and how it is structured, how to make the most efficient use of time limits, how to "attack" the different kinds of questions, and when an "educated" guess using partial knowledge is sensible. Students with such skills and knowledge about test taking are able to perform to the best of their ability. For this reason students should be encouraged to study the material in Taking the SAT carefully and to complete the sample test that is included. Schools may choose to assist students in the process through group meetings and discussion sessions to emphasize the importance of this preparation.
>
> Research continues to show that a review of mathematics concepts may be beneficial for students who are not enrolled in a mathematics course at the time the SAT is administered. Under such circumstances, the review supplements previous classroom instruction in mathematics. Some students do lead lives without significant exercise in complex mathematical thought, and a review of mathematics concepts may, therefore, serve to sharpen the mathematical reasoning ability of these students. Verbal reasoning, on the other hand, is an inescapable part of our daily lives; not surprisingly, verbal reasoning ability as measured by the SAT has not been shown to be as responsive to similar instructional efforts.[73]

Given the BCP's findings and the previous studies which did find significant positive effects from coaching, the analyses of coaching contained in the 1979-80 SAT materials did not appear to fully apprise students or educators of the possible benefits of coaching. The Commission shared

---

73 Educational Testing Service, ATP Guide for High Schools and Colleges 1979-81 13 (College Entrance Examination Board 1979) (citations omitted).

staff's concerns about the apparent inconsistency between the SAT materials and the research on coaching. Staff was authorized to bring these concerns directly to the attention of both ETS and the College Board and to offer the organizations an opportunity to respond.

Both organizations responded to the Commission's inquiry. ETS submitted a copy of a report on coaching now published in a booklet entitled The Effectiveness of Coaching for the SAT: Review and Reanalyis of Research from the Fifties to the FTC ("ETS Report").[74] College Board submitted a memo it had sent to the members of the College Board, as well as SAT descriptive materials for the 1980-81 testing year.

E. ETS Report

In addition to lengthy and detailed analysis of the BCP study, the ETS Report includes a review of studies regarding coaching for the SAT conducted before the BCP study.[75]

---

74 Messick, S., The Effectiveness of Coaching for the SAT: Review and Reanalysis of Research from the Fifties to the FTC (Educational Testing Service 1980) [hereinafter cited as ETS Report].

75 The review of coaching studies includes estimates of the "adjusted average score increases" for the Marron and Pallone studies. For the Pallone programs the estimates are 81 point increases on verbal for the 90 minutes per day six-week program and 68 points on verbal for the daily 50 minute, six-month program. For the Marron studies, the weighted average increases were 35 points on verbal and 54 points on math. ETS Report, supra n.74, at 10-13 and Table II-2 on 25.

We will not here review the entire 135 page report, but do wish to comment on a few of the issues raised by ETS regarding the BCP's analysis.

Many of ETS's comments focused on the fact that the BCP study was not based on an experimental design.[76] As noted supra, the BCP did not identify students and then randomly assign them to a group that would be coached and a group that would serve as a control. Rather, the BCP evaluated students who had voluntarily sought out commercial coaching and compared these with a control group. While ETS acknowledges that the BCP's analytical design did control for a number of background factors that could be associated with self-selection, ETS emphasizes that it is impossible to know and control for all background differences and all possible impacts of self-selection.[77]

ETS argues that the impact of self-selection is most likely to mean that the BCP's results overstate the effect of coaching, that whatever factors motivated students to take a coaching course make it likely that such students would have done better on the SAT even in the absence of coaching.[78] The Report does refer to the possibility that the

---

76 See, e.g., ETS Report, supra n.74, at 3 and 33.

77 ETS Report, supra n.74, at 32-33.

78 Id. at 61.

impact of self-selection causes not controlled for by either the BCP or ETS analyses could be to underestimate the actual effect.[79]

In considering the self-selection issue, it is important to remember that, in addition to having controlled for a variety of background factors such as grades and reported family income, the BCP examined the data for the possiblity of one likely reason that students would seek commercial coaching - whether the coached students had performed below their expectations on prior testing.[80] The concern was that the lower than expected prior test scores might be a statistical accident and that such students would thus perform better on subsequent testing regardless of coaching. The BCP's analysis did indicate that the coached students at the effective school were underachievers on standardized tests, but also found that those students would probably have continued to underachieve in the absence of coaching. Thus, the BCP concluded that its estimates of coaching effects, approximately 25 points on each of the two sections of the exam, did not have to be adjusted.[81]

---

79 Id. at 59.

80 BCP Report, supra n.36, at 24-34.

81 BCP Report, supra n.36, at 34.

After the general commentary on the BCP study, the ETS Report presents the results of two major statistical reanalyses undertaken by ETS. One, conducted by T.W.F. Stroud, used a statistical model similar to the BCP model but with certain additions and modifications, and also included data from a third coaching school which the BCP had not evaluated because of its small sample size.[82] Stroud also investigated the interaction between the size of the coaching effects at the two largest coaching schools and certain background data, such as race and income.[83]

The Stroud analysis confirms the BCP's estimate of effects. Stroud found what he termed "combined coaching/self-selection" effects of about 20-33 points for the mathematical section and 27-34 points for the verbal section for students at the effective school.[84] These results closely parallel the BCP's findings of an average coaching effect at that school of approximately 25 points on each of the two sections of the test. Stroud found negligible and inconsistent effects at the other two schools.[85] At the larger of those

---

82 ETS Report, supra n.74, at 3.

83 Id. at 3. The Stroud analysis is discussed at pages 43-51 of the ETS Report and set forth in Appendix 2.

84 Id. at 46.

85 Id. at 46.

two schools, Stroud did find interactions between both race and reported parental income and the "combined coaching/self-selection" effect on the verbal section of the exam.[86]

The second reanalysis, conducted by Donald Rock, applied a statistical model which is designed to take into account any differential rates of growth in SAT scores for the coached and control groups.[87] He identified groups of coached (at the effective school) and control students for whom scores were available on the Preliminary Scholastic Aptitude Test ("PSAT") and two administrations of the SAT.[88] The growth rates beween the PSAT and the first SAT were determined for both groups. Rock found that the coached and control groups were growing at virtually the same rate on the mathematical section of the tests, but that the coached group was growing faster than the control group regarding the verbal section.[89] Rock assumed that these growth rates, absent coaching, would continue unchanged between the first and second administration of the SAT.[90] The

---

86 Id. at 47-51.

87 Id. at 3-4. The Rock analysis is discussed at pages 53-59 of the ETS Report and set forth in Appendix 3.

88 Id. at 54.

89 Id. at 55-57.

90 Id. at 57.

growth model was then used to estimate the coaching effect while controlling for the difference in growth rates. Rock found that if one assumed that the coaching took place, on average, midway between the first and second administrations of the SAT, the effects were approximately 30 points on the mathematical scores and 11 points on the verbal scores.[91] However, if one assumed that the coaching took place immediately after the first administration of the SAT, the effect was about 31 points on mathematical scores and 17 points on verbal.[92]

Rock noted that it was more likely for students to obtain coaching as close to the second administration of the SAT as possible, and thus concluded that the 11 point figure was probably more accurate than the 17 point figure in assessing the verbal effect.[93] In discussing the Rock Analysis in the Overview of the ETS Report and in a paper entitled "The Effectiveness of Coaching for the SAT: Review and Reanalysis of Reseach from the Fifties to the FTC - Summary of Issues and Results" disseminated at the end of May, ETS discussed only Rock's 11 point finding regarding

---

91 Id. at 57-58.

92 Id. at 57-58.

93 Id. at 58, 132.

the verbal section of the test.[94] The assumption about the timing of the coaching seems reasonable since students are likely to believe that the closer the preparation is to the test in question, the more likely one is able to best utilize the training. But, the assumption makes it difficult to understand why the effect was found to be smaller if the coaching took place closer in time to the test in question.

Substantial efforts were made by FTC staff and consultants to replicate the computations that led to Rock's 11 point finding. Because these efforts were unsuccessful, the BRO requested an explanation of and support for the 11 point finding. Shortly after the BRO inquiry, ETS published an Errata Sheet concerning the Rock analysis. In the Errata, ETS abandoned its reliance on the 11 point finding and adopted instead the higher 17 point finding for the verbal section of the test. Because the FTC's efforts to replicate the 11 point finding were unsuccessful and publication of the Errata came shortly after the BRO inquiry, Commission staff concludes that the 11 point finding was due to a conceptual or computational error.

---

94 Id. at 4; S. Messick, "The Effectiveness of Coaching for the SAT: Review and Reanalysis of Research from the Fifties to the FTC - Summary of Issues and Results" 5-6 (Undated).

In considering even the 17 point finding, at least one observation should be kept in mind. The Rock analysis is based on the assumption that the coached group, who did start out with higher scores than the control group, would have continued to grow at the same faster pace absent coaching. While the assumption of linearity is not unreasonable, it is impossible to verify its accuracy.[95] It is at least possible that in the absence of coaching, the coached groups' learning would have peaked or at least tapered off between the first and second administrations of the SAT, in which case Rock's 17 point finding underestimates the coaching effect. Finally, regardless of whether the verbal effect was the 17 points found by Rock, the 27-34 points found by Stroud or the approximately 25 points found by the BCP, the effects estimated through all the analyses are significant.

F. SAT Descriptive Materials for 1980-81

As noted earlier, the College Board provided the Commission with copies of the new descriptive materials on the SAT for 1980-81. Of most significance is a new statement on coaching, distributed to secondary schools this fall, entitled "Six Points about Special Preparation for the SAT. A Message

---

95 Other researchers have commented on possible problems presented by the linear growth model. A. Bryk, J. Strenio & H. Weisberg, "A Method for Estimating Treatment Effects When Individuals are Growing", 5 J. of Educational Statistics 5, at 25-28 (1980).

to Students" ("Six Points Message"). This message is a much more comprehensive statement on coaching than has been provided to students in the past. It makes clear that there are different types of special preparation offered in different settings and acknowledges that different results may occur. The message includes the following discussion of studies on the effects of coaching.

> Studies of special preparation programs carried on in many high schools show various results averaging about 10 points for the verbal section and 15 points for the mathematical over and above the average increases that would otherwise be expected from intellectual growth and practice. In other programs results have ranged from virtually no improvement in scores to average gains as high as 25-30 points for particular groups of students or particular programs. Recent studies of commercial coaching have shown a similar range of results.

Thus, while the message still presents the average of effects found in studies of some number of high school programs, it does go on to note that studies of other programs have shown greater increases. Perhaps most important is the fact that the message acknowledges that improvements as high as 25-30 points on a section of the SAT are possible. While the statement may not reflect the very upper limits found in studies such as Pallone, it is far superior in its breadth and detail than earlier statements made in the SAT descriptive materials.

Unfortunately, statements regarding coaching in the 1980-81 versions of neither the bulletin nor guide for students

nor the guide for counselors and admissions officers parallel the Six Points Message. These materials appear to lump previous studies together to come up with a 10 point coaching effect for verbal and 15 points for math.[96] Even if these numbers were produced by a fairly computed averaging of all prior studies, the number tells students and educators little. Students who wish to properly prepare themselves for college admissions are not interested in such statistical averages. While they have a need to know that some forms of special preparation have been shown to be of little effect, they have a right to know that other studies have shown significant positive improvements. It is essential that students be provided with enough information to make

---

96 Educational Testing Service, Student Bulletin 1980-1981, at 12 (College Entrance Examination Board 1980); Educational Testing Service, Taking the SAT 3 (College Entrance Examination Board 1979); Educational Testing Service, ATP Guide for High Schools and Colleges 1979-81, at 13 (College Entrance Examination Board 1980). The 10 point verbal effect and 15 point math effect may be based on only certain coaching studies. For in summarizing prior research, the ETS Report notes that "the average coaching effect across studies having some type of control group was less than 10 points for SAT-Verbal and less than 15 points for SAT-Math." ETS Report, supra n.74, at 29. These estimates do not include the BCP findings or the findings of studies which did not have control groups. With respect to the latter, the ETS Report states: "The provisional estimate of average program effects for these noncontrolled studies was 38 points for Verbal and 54 points for Math." ETS Report, supra n.74, at 29.

an intelligent decision about whether to seek special preparation. Similarly, but perhaps even more importantly, it is essential that educators who rely on the test to advise and to make admissions decisions be fully informed about this important topic.

By letter to staff of the Commission, College Board has indicated that it intends to revise the bulletin for students and other appropriate SAT materials for the 1981-82 testing year to include the Six Points Message or its equivalent. It is hoped that in making future revisions the College Board continues on the course toward full information begun by the Six Points Message.

V. Implications of the Federal Trade Commission Findings on Coaching

In the course of the investigation, especially in light of the results of the BCP study, it became clear that many educational policy questions are raised by a finding that coaching for a standardized admission test such as the SAT can be effective. Perhaps the most fundamental question is simply "Why?" The research conducted to date does not provide an answer to the question. The focus of coaching studies has been on whether coaching is effective rather than why.

In its recent report, ETS recognizes that a finding that coaching could substantially improve scores would

have important educational implications.[97] ETS also recognizes the question of "why" coaching might work as an important one.

> One key issue is the extent to which increased test scores attributable to coaching may represent stable long-term improvements in the verbal and mathematical reasoning skills measured by the SAT or instead reflect the overcoming of inadvertent sources of test difficulty unrelated to these reasoning skills, such as difficulty associated with test anxiety and unfamiliarity with different item formats and test-taking strategies.[98]

As ETS sees the issue, the question is whether the score improvements relate to the substance of the test or the test's construction.

Lewis Pike, in the 1978 literature review discussed supra at pages 22-25, notes in somewhat more detail the possible ways that coaching could affect scores.[99] Through instruction regarding the basic skills that the SAT is designed to measure, the student could be better prepared to demonstrate the underlying knowledge and reasoning abilities he/she has already acquired during years of formal schooling.[100] Though students may have a wide base of knowledge of those skills being tested, coaching might help to improve scores

---

97 ETS Report, supra n.74, at 4, 61-67.

98 Id. at 4.

99 Literature Review, supra n.52, at 71.

100 Id.

by filling in any gaps between that underlying base and the actual test content.[101] Instruction in testwiseness skills could help the students achieve a better score by familiarizing them with the item formats, and by explaining test strategies, such as when to guess on the basis of limited information.[102] Coaching could help by reducing students' anxiety level or by helping them to cope efficiently with a timed test format.[103]

At least one other explanation for possible score improvements is that coaching could teach students how to answer questions on the test without any understanding of the skills which are being tested. For example, a test might contain extraneous clues that the student could use to select the correct answer. However, as ETS notes in its report, in a well designed test such extraneous clues should be rare.[104]

The answer to why coaching may be effective is important in determining the implications of coaching for test validity. First, what are the implications for validity if coaching is improving the skills the SAT is designed to measure? In ETS's

---

101 Id.

102 Id.

103 Id.

104 ETS Report, supra n.74, at 65.

view, if any improvement in scores is related to long-term improvements in the underlying skills being measured, then the possibility of such improvements does not invalidate the SAT.[105] But, such improvements may say something about the skills being measured by the test. The authors of the 1972 coaching study discussed supra at pages 21 and 22 addressed this issue. "[I]f special instruction were found to influence SAT scores substantially, the validity of the test would be open to question, since it is intended to be a measure of relatively stable attributes developed over a long period of time."[106]

Skill related coaching which affects something other than long-term improvements should also be considered. If coaching is serving a review function, then a well reviewed student's score is likely to be a better indicator of the student's actual potential.[107] On the other hand, if the student is learning the skills being tested, but retains them only long enough to perform well on the test, the test's ability to predict future academic performance may be reduced.

Also to be considered are implications for test validity if coaching can help students to improve scores

---

105 Id. at 4, 62-63.

106 L. Pike & F. Evans, Effects of Special Instruction for Three Kinds of Mathematics Aptitude Items 4 (College Entrance Examination Board 1972) (citations omitted).

107 Literature Review, supra n.52, at 71.

without a concomitant improvement in underlying skills. Alderman and Powers noted in their 1979 review of verbal coaching programs two possible implications.

> To the extent that special preparation for an aptitude test leads to score changes unrelated to the criterion (e.g., college grade point average or job performance), that preparation actually impairs the test's usefulness as a predictor of later success. On the other hand, when a test contains unusual or unfamiliar item types, special preparation may help students in attaining scores more indicative of their ability.[108]

The question then is whether instruction in testwiseness is permitting the students to somehow obtain inflated scores, or simply to better demonstrate their underlying abilities. In the first case, validity is questioned.[109] In the second case, at least for those students' scores, it is not.[110]

In assessing the question of test validity, however, it is important to consider not only the validity of the

---

108 D. Alderman, & D. Powers, The Effects of Special Preparation on SAT-Verbal Scores 1-2 (College Entrance Examination Board Research and Development Reports, 78-79, No. 4, February 1979).

109 The ETS Report notes that "if improved test wiseness leads to test scores that are inaccurately high, the interpretive validity of the test would be diluted and its predictive validity jeopardized." ETS Report, supra n.74, at 63.

110 The ETS Report notes that: "If improved test wiseness increases test scores that were inaccurately low because of construct-irrelevant difficulty, a more accurate assessment of ability level would result and the predictive value of the test should be enhanced." ETS Report, supra n.74, at 63.

coached student's score, but also the validity of comparing the scores of different students. Is it appropriate to compare the score of a coached student with the score of an uncoached student? The question is particularly important because the materals reviewed during the FTC's investigation indicate that all students do not have equal access to coaching.

As noted supra, tuition costs at commercial coaching schools can be high. With costs of up to $300 for an SAT course, many students are clearly priced out of that market. The BCP's analysis of the coached and uncoached students included in its study indicates that it is the more affluent student that is likely to be coached at commercial schools. The BCP found a statistically significant difference in the parental incomes of the coached and uncoached student groups it evaluated.[111] Over forty percent of the coached students reported parental incomes of $30,000 or more, while the comparable figure for the uncoached students was only slightly over seventeen percent.[112] Further evidence of the likely socio-economic split between coached and uncoached SAT test takers is found in the BCP's analysis of the type of high schools the students attended. Approximately

---

111 BCP Report, supra n.36, at 8.

112 Id.

seventy-five percent of the uncoached students attended public schools, while only about fifty-five percent of the coached group were public school students.[113]

We also suspect that secondary schools with a small percentage of college bound students are less likely to have the resources to offer test preparation courses to their students. Wealthier high schools may have the capacity to offer precisely the type of intensive review that has been shown to be effective in raising scores, while financially pressed schools may not.

The question of access takes on added significance if whatever benefits coaching confers are dependent on the characteristics of the students themselves. For it is at least possible that the students most in need of training may have the least access to it. For example, if instruction in testwiseness skills is effective, the students most in need of such instruction may be those who have had the least exposure to multiple-choice format examinations prior to taking the SAT. Students at schools in wealthy neighborhoods are more likely than their inner city school counterparts to have been exposed to a number of multiple-choice standardized tests before taking the SAT.[114]

---

113 Id. at 9.

114 Hearings on Testing, supra n.7, at 687 (Testimony
(Footnote Continued)

The Stroud analysis findings of an interaction between reported parental income and the "combined coaching/self selection" effect on the verbal section of the SAT also deserve further inquiry. The ETS Report suggests two possible explanations for why students reporting lower parental income experienced greater score increases after attending a coaching program than did those reporting higher parental income. One of the explanations suggests that economically disadvantaged students have the most to gain from such coaching.

> Since students with low family incomes may have had less access to learning resources at either home or school, this interaction might have arisen because some of them chose to attend coaching school as a limited-cost effort to gain short-term intensive access to such resources, which they then used to good compensatory effect.[115]

The rival hypothesis referred to the self-selection issues, that motivation was responsible for the score increases rather than the coaching program itself.

> [S]ince both motivation and financial means are probably instrumental in deciding whether or not to attend a commercial coaching school, it is not unreasonable that coaching school enrollees having lower than average

---

114 (Footnote Continued)

of Althea Simmons, Director, Washington Bureau of the National Association for the Advancement of Colored People).

115 ETS Report, supra n.74, at 48.

financial means might have especially strong motivation to perform well.[116]

Stroud's finding of an interaction between race and the "combined coaching/self-selection" effect on verbal scores, where the "coaching/self-selection" effect was larger for blacks than for nonblacks, also raises interesting questions about differential coaching impacts.[117] While the sample size evaluated was very small, the report notes that because of the size of the effect and its statistical significance, further research is clearly necessary.[118]

In a recent presentation at an American Psychological Association symposium, Lewis Pike presented his analysis of the equity of access issue.[119] He noted that when coaching works, and there is unequal access, a profound unfairness is done to the students without access. He also noted that for students who were coming from already disadvantaged educational backgrounds, any lack of preparation in how to cope with a test constituted a double jeopardy.

---

116 Id.

117 Id. at 3, 48-51.

118 Id. at 51.

119 Paper presented on September 1, 1980, in Montreal, Canada, at the American Psychological Association Symposium on Truth in Testing Legislation. Mr. Pike, who had served as an ETS researcher for a number of years, is now with the National Institute of Education of the Department of Education.

ETS has also commented on the equity of access issue, making some distinction on the basis of what the coaching program is accomplishing. If a coaching program could improve underlying skills, then the issue of equity of access "is thus similar to the issue of equity of access to effective school programs or effective life experiences. Thus, coaching or special preparation programs producing increased test scores by improving the abilities measured would have important implications for educational practice and social policy."[120] If a coaching program could be effective because of the testwiseness training given, then the issues of equity of access are also important, but the implications revolve around testing practice. ETS would then recommend that any difficulty related to the format of the test be reduced and that substantial efforts be made to increase test familiarization and testwiseness.[121]

The final issue then, which ETS's comments have raised, is what steps should be taken if coaching is effective but access is not open to all students. Before pursuing the options that ETS has suggested, or others, it is first important to fully understand the precise skills and knowledge that the SAT is in fact testing, to evaluate the value

---

120 ETS Report, supra n.74, at 64-65.

121 Id. at 65-66.

of developing such skills and knowledge, and to evaluate the value of training students to demonstrate those skills on any particular examination.[122] Further research and analysis are clearly necessary before answers can be found to the many important educational questions raised by a finding that coaching can be effective.

VI. Conclusion

The purpose of the FTC's investigation has been to assure that the various organizations involved in the standardized testing industry accurately explain the effectiveness of coaching. Progress has been made and staff is continuing to monitor the market to see that full and fair disclosure becomes the norm.

In the course of its investigation, the FTC staff has produced an important study regarding the effectiveness of commercial coaching for the SAT. However, the investigation could not and has not attempted to resolve the many educational

---

122 These questions implicitly include questions regarding the predictive validity of the test noted supra at n.6 and the possibility of racial, cultural or economic bias. See, e.g., Hearings on Testing, supra n.7, at 545-55 (Statements of James Loewen, Director, Center for National Policy Review, School of Law, Catholic University of America). Further, the value of multiple-choice aptitude and achievement tests has been criticized because such tests do not measure motivation or creativity, and concern raised about whether they properly recognize the talents of the deep and subtle thinker. B. Hoffman, The Tyranny of Testing 91-101 (1962).

policy questions raised by a finding that coaching for a standardized admission test such as the SAT can be effective. Researchers have only begun to address these questions. We have discussed some of them here to encourage experts involved in the field to invest their expertise and resources to find appropriate answers.